

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: June 18, 2003, 15:04:29 / Search time 16.0187 Seconds  
(without alignments)  
2160.495 Million cell updates/sec

Title: US-09-807-933B-5

Perfect score: 1956

Sequence: 1 MKPLTIASAILALAVGTEM.....TYKQVTCPKATKSGCSRK 360

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	702	35.9	229	2 JC7308	cellulase (EC 3.2.
2	524.5	25.8	511	2 S10527	endoglucanase B pr
3	490	25.1	393	2 S59499	cellulase eg11 - s
4	226	11.6	471	1 A26160	cellulose 1,4-beta
5	226	11.6	471	1 A38979	cellulose 1,4-beta
6	198.5	10.1	475	2 S49886	probable membrane
7	196	10.0	662	2 A45155	mucin PIM-C.1 - Af
8	179.5	9.2	888	2 T46726	secreted acid phos
9	173	8.8	328	2 S67570	hypothetical prote
10	173	8.8	1217	2 S52714	sericinB - silkw
11	168.5	8.6	371	2 S20075	promastigote surfa
12	168	8.6	1832	2 T31113	mucin-like glycop
13	164	8.4	797	1 VGBEX1	glycoprotein X pr
14	164	8.4	2761	2 T21064	hypothetical prote
15	163.5	8.4	402	2 E86185	hypothetical prote
16	163.5	8.4	605	2 T33913	hypothetical prote
17	162.5	8.3	418	1 S28372	cellulase (EC 3.2.
18	162.5	8.3	1367	2 S51959	hypothetical prote
19	162	8.3	796	2 T21460	hypothetical prote
20	161	8.2	477	2 S53362	mucin 5AC (clone J
21	160.5	8.2	786	1 A47547	serine proteinase
22	160	8.2	327	2 S20074	promastigote surfa
23	160	8.2	503	2 S63257	probable membrane
24	160	8.2	1275	2 T33369	hypothetical prote
25	158	8.1	354	2 T46740	microfilament shea
26	158	8.1	1161	2 S57180	probable membrane
27	157.5	8.1	344	2 T40167	hypothetical prote
28	156.5	8.0	263	2 S01360	salivary glue prot
29	156.5	8.0	750	2 T42614	probable envelope

30	156	8.0	438	1 S70602	cellulose 1,4-beta
31	155.5	7.9	785	2 T16509	hypothetical prote
32	155.5	7.9	1251	2 T21389	hypothetical prote
33	155	7.9	217	2 S01358	salivary glue prot
34	154.5	7.9	1032	2 T34433	hypothetical prote
35	154.5	7.9	2232	2 T34434	hypothetical prote
36	153.5	7.8	725	2 A41258	a-egg-luciferin core
37	153	7.8	299	2 S56031	pathogenesis-relat
38	153	7.8	1360	2 T33922	hypothetical prote
39	152.5	7.8	798	2 T34248	hypothetical prote
40	151.5	7.7	971	2 B50835	probable tail fibre
41	151.5	7.7	973	2 C85693	probable membrane
42	151	7.7	1566	2 T20058	hypothetical prote
43	150	7.7	770	2 T22808	hypothetical prote
44	149	7.6	605	2 S48940	hypothetical prote
45	148	7.6	446	2 T45525	Wsc4 homolog (limp

## ALIGNMENTS

### RESULT 1

UC7308 cellulase (EC 3.2.1.4) - Scopulariopsis brevicaulis

N:Alternate names: endoglucanase I

C:Species: Scopulariopsis brevicaulis

C>Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000

C/Accession: UC7308; PC7087

R:Nakatani, F.; Kawaguchi, T.; Takada, G.; Sumitani, J.; Moriyama, Y.; Arai, M.

Biosci. Biotechnol. Biochem. 64, 1238-1246, 2000

A:Title: Cloning and sequencing of an endoglucanase gene from Scopulariopsis brevicaulis

A:Reference number: UC7308

A/Accession: UC7308

A:Molecule type: DNA

A:Residues: 1-229 <NAK>

A:Experimental source: strain TOF-1212

A/Accession: PC7087

A:Molecule type: protein

A:Residues: 21-37/149-164 <NA2>

C/Genetics:

A:Gene: eg1

A:Introns: 147/3

C/Keywords: glycosidase; hydrolase

Query Match 35.9% Score 702; DB 2; Length 229;

Best Local Similarity 56.4% Pred. No. 1.4e-36;

Matches 119; Conservative 34; Mismatches 52; Indels 6; Gaps 4;

QY	152	ASGNGETTRVWDCKKPCSCWPGKADVT-SPVGSCKNDKTLADNN-TONGC-VGGSSYTC	208
DB	21	ASGCTTTRVWDCKKPCSCWPGKADVT-SPVGSCKNDKTLADNN-TONGC-VGGSSYTC	80
QY	209	NDNPNVWVDDIAYGFAAASISGSEATWCACFELFTSTAVYKXQVNTVNGSDLG	268
DB	81	SSHPWAVDDEIAYGMAVNIIGQTESDWCACFELFTSTAVYKXQVNTVNGSDLG	140
QY	269	SNTGAPFLQMPGGGCVGYNCATQWCAPTGKRGARVGSASADCSULPSALQAGCKWR	328
DB	141	NN--HFDIAMPGGGCVGYNCATQWCAPTGKRGARVGSASADCSULPSALQAGCKWR	197
QY	329	PGWFPNDNPMTVTKQVTCPKATKSGCSR	359
DB	198	PDWFGTDNPVTSFREVCEPALVQSGCOR	228

### RESULT 2

S10527 endoglucanase B precursor - Pseudomonas fluorescens

C:Species: Pseudomonas fluorescens

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 08-Oct-1999

C/Accession: S10527

R:Gilbert, H.J.; Hall, J.; Hazlewood, G.P.; Ferreira, L.M.A.

Mol. Microbiol. 4, 759-767, 1990

A>Title: The N-terminal region of an endoglucanase from *Pseudomonas fluorescens* subspect  
A:Reference number: S10527; MUID:90355836; PMID:2117693  
A:Accession: S10527  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-511 <GIL>  
A:Cross-references: EMBL:X52615; NID:g45497; PID:CAA36844.1; PID:g45498  
C:Superfamily: bacterial cellulose-binding domain homology; glycosidase GWGW domain hom  
F:31-128/Domain: bacterial cellulose-binding domain homology <BCB>  
F:180-217/Domain: glycosidase GWGW domain homology <MGW>  
F:32-127/Disulfide bonds: #status predicted

Query Match 26.8%; Score 524.5; DB 2; Length 511;  
Best Local Similarity 30.8%; Pred. No. 2.6e-25;  
Matches 123; Conservative 59; Mismatches 119; Indels 99; Gaps 12;

QY 33 CGGKMDGPTCCSGSTCVDPDNPYSCVNPENLTGNSKHTTTEASKTTTG 92  
DB 127 CGGQSSABSVASSSS-----SSSVSSSTPRSSSSSVSPGTSSTSS 171  
QY 93 SKTTTTEASK-----KTT-----TTEASKTTT 115  
DB 172 SSVLGAGACNNVGTLPLONNNSNGWYEDGRSCVARTTCAQAPYGVIVTSSSTPS 231  
QY 116 TTEASKTTTTHKASTSTSSSSASTYSAVSGASNGETTRWDCCKPSCMPGKA 175  
DB 232 SSSSSSSSVASSSSSSSATSSASVS-SVPPIDGCG-NGVATRWDCCKHCGMSANV 288  
QY 176 -VTSFVSGCNKDGKTLADNNTQNGCVGSSYTCNDNOPWVSDLAFAAASISGSE 234  
DB 289 PSLVPSLOCSANNRLSDVSVSGSCDGGGYCMCKIPFVSPILAVGYAITS-SGDV- 346  
QY 235 ATWCAACFELTFT-----STAVKGMVVOVNTGSDLGSGNTGHPDLOMPGG 283  
DB 347 ---CGRCYQLOFTGSSYNAPGDPGSAALAKMTIVATNIGDV---SGCGDILVPGG 400  
QY 284 VGTNGCATONGAPDGMARVGG-----VSASDC-----SNLPS 319  
DB 401 VAFAPACSHQMGVSNALGAQYGGFLAACKQOLGYNASISQYKCVLNCDSVFSRGIT 460  
QY 320 ALQAGCKRFPKFNADNPMTYKQVTCFKAITAKSGCSR 359  
DB 461 QLOGCTFAEMFPAADNPDLKKEVPCPAELITTSGMNR 500

RESULT 3  
SS9499  
cellulase egII - smut fungus (Ustilago maydis)  
C:Species: Ustilago maydis (corn smut)  
C>Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 28-May-1999  
C:Accession: S59499  
R:Schaewecker, F.; Wanner, G.; Kahmann, R.  
Biol. Chem. Hoppe-Seyler 376, 617-625, 1995  
A>Title: Filament-specific expression of a cellulase gene in the dimorphic fungus *Ustilago*  
A:Reference number: S59499; MUID:96145728; PMID:8590631  
A:Accession: S59499  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-393 <SCH>  
A:Cross-references: GB:S81598; NID:g1478378; PID:AA836147.1; PID:g1478379

Query Match 25.1%; Score 490; DB 2; Length 393;  
Best Local Similarity 46.1%; Pred. No. 2.7e-23;  
Matches 101; Conservative 29; Mismatches 67; Indels 22; Gaps 9;

QY 156 GETTYWDCCKPSCMPGKADVTSPVSGCNKDGKTLADN---NONGCVGSSSTQNDN 211  
DB 27 GNAITRWDCCLASASWEGAPVAPVADKADGVTLLDSKDPSSQSGCNKCKMPCSM 86  
QY 212 QPWWVSD--LAVGFALASISGSEATWCCAFELTF---TSTVVKGMVVOVNTGSS 265  
DB 87 QPFDDETPTLAFGCA--FTTGQBSDDTCACFYAEFHDAGCKMKRKLIFQVTTNGG 144

QY 266 DLGNTGAHFDLQMPGGVGIY-NCATOWGAPTDGMGARVGVSSASDCSNLPSALOAG 324  
DB 145 DVQSQ---NFPBQIFGGGLGAFPKGCPAOWGVBSLWMDQVGVKASATCKLPRLDQG 201  
QY 325 CKMRFG-WFKANDPTM--TYKQVTCPAITAKSGCSCK 360  
DB 202 CKMRSEW---GDNPVLKGSFPRVPCPKSLIDRSQCCK 237

RESULT 4  
A26160  
cellulose 1,4-beta-cellubiosidase (EC 3.2.1.91) II precursor - fungus (*Trichoderma reesei*)  
N:Alternate names: exo-cellubiohydrolase II  
C:Species: *Trichoderma reesei*  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A26160; A26472  
R:Chen, C.M.; Gitzals, M.; Stafford, D.W.  
Bio/Technology 5, 274-278, 1987  
A>Title: Nucleotide sequence and deduced primary structure of cellubiohydrolase II from  
A:Reference number: A26160  
A:Accession: A26160  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-471 <CH>  
R:Teeri, T.T.; Lehto-Vaara, P.; Kauppinen, S.; Salovuori, I.; Knowles, J.  
Gene 51, 43-52, 1987  
A>Title: Homologous domains in *Trichoderma reesei* cellubiohydrolase enzymes: gene sequence an  
A:Reference number: A26472; MUID:87248061; PMID:3596237  
A:Accession: A26472  
A:Molecule type: DNA  
A:Residues: 1-471 <TE>  
A:Cross-references: GB:M16190; NID:g170540; PID:AAA34210.1; PID:g170541  
C:Genetics:  
A:Insertions: 31/2; 160/2; 243/1  
C:Superfamily: cellulose 1,4-beta-cellubiosidase II; fungal cellulose-binding domain hom  
C:Keywords: glycosidase; hydrolase; polysaccharide degradation  
F:31-62/Domain: fungal cellulose-binding domain homology <FCB>

Query Match 11.6%; Score 226; DB 1; Length 471;  
Best Local Similarity 25.5%; Pred. No. 5.9e-07;  
Matches 94; Conservative 41; Mismatches 156; Indels 78; Gaps 13;

QY 5 TIASAILALAVGTEMAHAECSKAYVQCGKNDGPTCCSGSTCVDPDNPYSCV 64  
DB 8 TLATLATAASVPLEROA---CSSVWCGCGGQNMSPCCASGSTCV--YENDVYSGCLP 63  
QY 65 NENLSTNKSHTTTTESAKTTTGSKTTTTEASKTTTBASKTTTTEASKTT 124  
DB 64 G---AASSSTRAASTTRVSPPTTSRS--SATPPGSTTRVPVPGSGTATYSGNPEV 118  
QY 125 TTYKASTSTSSSSSSAS---TVYSAVSGASGNGETTRWDCCKPSCMPGKADVTSP 180  
DB 119 GVTPPANNVYASVSLAIPSLTGAMTAABAIVK-----VPSFMWLDLDTKPL 168  
QY 181 VGSCKDGKTLADNNTQNGCVGSS---YTCNDNOPWVSDLAFAAASISGSE 235  
DB 169 M-----EQLTADIRTNKNGNVAQGFVVYDLPDRCALASNGEY-----SIADGVA 217  
QY 236 TW-----CAACFELTFTSTAVKGMVVOVNTGSDLGSGNTGHPDLOMPGGGCI 286  
DB 218 KYKVIIDTIRQIVVEYSIRTLVIEPDSLNVLTNLTGPRCANAGSAV-----DEC 269  
QY 287 YNGCATOWGAPTDGMGARVGVSSASDCSNLPSALOAGCKRFPKFNADNPMTM---TY 342  
DB 270 INVAVTQNLN-----NVAMWLDLGHGMLGMPANQDPAQLFANVY 311  
QY 343 KQVTCFKAI 351  
DB 312 KNASPRAL 320

RESULT 5  
A38979





```
Qy      301 WG--ARYGVSSASDCSN 316
          :  :| | | |
Db      956 YSSDSRDGVSSTGSSN 973
```

## RESULT 11

pmaestgroc surface antigen p2 (clone 2.5) precursor - Leishmania major (fragment)  
 C:Species: Leishmania major  
 C>Date: 13-Jan-1995 #sequence\_revision 06-Feb-1998 #text\_change 31-Jan-2000  
 C:Accession: S20075; C41710  
 R:Murphy, P.J.; Spithill, T.W.  
 J. Biol. Chem. 266, 24477-24484, 1991  
 A:Title: Variants of a Leishmania surface antigen derived from a multigenic family  
 A:Reference number: A41710; MUID:92105105; PMID:1761547  
 A:Source: GenBank

Query Match	8.6%	Score	168.5	DB	2	Length	371
Best Local Similarity	27.8%	Pred.	No. 0.0017				
Matches	58	Conservative	28	Mismatches	100	Indels	23
						Gaps	6

```

QY      47  GATCVDYDNPVSGCVPEMNLSTNNKSHKTTTTEASAKTTTTKSKKTTTTEASKTT 106
Db      162  GSDCLAGADCTTTTKKLPTTTTTTTKKPPTTTTTTTKKPPTTTTTTTKKPPTTTTTTTKKPPTT 221
QY      107  TTEASKTTTTEASKK---TTTTTKKASTSTSSSSASTNY---SAVSGASGANGET 158
Db      222  TTTTTPPTTTTTTKKPPTTTTTTTKKPLLTATTKKPPTTTTTTTKKPPTIITSTTKLPPTT 281
QY      159  TRYNDCCKPSGCMFGKADYTSFVSGCNKDGKTLADNNTONGCVGSSSYTC-NDNQPVAVS 217
Db      282  TTEAPAEPTTTATPNTPTPAPETECEVDGCEV---CEGDAAACARSGREDYFILT 313
QY      218  DD---LAV---GFAAASISGSGEATWCCA 240
Db      334  DERTCLVYCDGGVAAVSSGVAATAAVCA 362

```

## RESULT 12

mucin-like glycoprotein 900 - Cryptosporidium parvum  
 C:Species: Cryptosporidium parvum  
 C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999  
 C:Accession: J31113  
 R:Barner, D.A.; Bonnin, A.; Huang, J.X.; Goussset, L.; Wu, J.; Gut, J.; Doyle, P.; Dubremet, M.L. Blochom. Parasitol. 96, 93-110, 1998  
 A:Title: A novel multi-domain mucin-like glycoprotein of Cryptosporidium parvum mediates adherence to host cells  
 A:Reference numbers: 220989; PMID:9506935; PMID:9851610  
 A:Accession: J31113  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-1832 <BAR>  
 A:Cross-references: EMBL:AF068065; NID:g4063041; PTD:g4C63042; PTDN:ACG98153.1

Query Match	8.6%;	Score 168;	DB 2;	Length 1832;
Best Local Similarity	24.3%;	Pred. No. 0.008;		
Matches 77;	Conservative	38;	Mismatches 108;	Indels 94;
				Gaps 13

[illegible]

```

Qy 180 -----PVGSG-C--NNDGGTLADNNTQNGCGV-----GSSYTCNDNQPNVWVSDDLAYGFAA 226
Db 524 SVTKPEBMCLEKNGECBAKATTVGVIGKGRILENGAFI-----MIPND-----57070
Qy 227 ASISGSGSEATWCCACFELTFTSTAVKGGKMMVQVNTNGS-DLGSNTG-----AHFDLQM 2797
Db 571 -----THRFRPKVMDVNTISVRCGKAGLTFPDRSLPTI 608
Qy 280 PGGGVGITNGCATQMGAPFDWGGAIRYGVSSASDCSNLPSALOAGCKKRFQMFKNADNPT 3339
Db 609 P--PVGHNHSCSLIIV--VSGGGKIHVSPYSGKDVLSAPIQ-----PC 6494
Qy 340 MTKQVTCPEKATPAKSG 356
Db 650 ELENEYTC--DTCIAKYG 665

```

### RESULT 13

A: glycoprotein X precursor - equine herpesvirus 1 (strain Ab4p)  
C: Species: equine herpesvirus 1  
A: Note: host Equus caballus (domestic horse)  
C: Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 16-Jul-1999  
C: Accession: H36802  
R: Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.  
S: Submitted to GenBank, March 1992  
A: Description: The DNA sequence of equine herpesvirus-1.  
A: Reference number: A36805  
A: Accession: H36802  
A: Molecule type: DNA  
A: Residues: 1-797 <TEL>  
A: Cross-references: GB:86664; NID: g330791; PIDN: AAB02506.1; PID: g330862  
R: Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.  
V: Virology 189, 304-316, 1992  
A: Title: The DNA sequence of equine herpesvirus-1.  
A: Reference number: A41831; MUID: 92295566; PMID: 1318606  
A: Contents: annotation; possible protein-coding frames  
A: Note: neither amino acid nor nucleotide sequence is given  
C: Genetics:  
A: Gene: 71  
C: Superfamily: equine herpesvirus glycoprotein X; equine herpesvirus 1 glycoprotein homologous to glycoprotein X of other herpesviruses  
C: Keywords: glycoprotein; transmembrane protein  
F: 1-22/Domain: signal sequence #status predicted <SIG>  
F: 23-797/Product: glycoprotein X #status predicted <MOT>  
F: 23-465/Region: serine/threonine-rich  
F: 489-797/Domain: equine herpesvirus 1 glycoprotein homology <EHG>  
F: 766-790/Domain: transmembrane #status predicted <TMH>  
F: 550/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match	8.4%	Score 164	DB 1	Length 797	
Best Local Similarity	24.2%	Pred. NC. 0.0064			
Matches 64	Conservative 36	Mismatches 127	Indels 38	Gaps 6	

[illegible]

